

## 動詞と主体の属性を用いた複文の接続関係の解析

向 伸      景頁<sup>†</sup>

本論文は、動詞と主体の属性を用いて、複文中の接続関係を解析するモデルを作成し、評価した結果を述べる。複文中の接続関係の关系的意味は、接続詞、助詞等の接続の表現だけでは決まらず、曖昧性がある。例えば、助詞「て」による接続関係には、「時間的継起」のほかに「方法」、「付帯状態」、「理由」、「目的」、「並列」などがある。これらの関係的意味は、従属節や主節の述語の表している事象の意味タイプ、およびその組み合わせによって決まってくる。従って、動詞と名詞の意味的關係を表すために、動詞と名詞の意味分類を用いた格パターンがあると同様に、従属節と主節の接続関係にも、各々の節を構成する動詞と主体の属性を用いた接続関係パターンが存在すると考えることができる。本論文のモデルでは、従属節と主節の、動詞と主体の属性を用いて、接続関係の关系的意味を推定する。動詞の属性として、意志性、意味分類、慣用的表現、ムード・アスペクト・ヴォイス、主体の属性として、主節と従属節の主体が同一かどうか、無生物主体かどうかを用いた。このモデルを、技術文書に適用した結果、95%の文が正しく解析できた。

キーワード： 接続関係、多義性解消、動詞の属性、主体の属性、接続関係パターン

## Analysis of Coherence Relation for Complex Sentences Using Attributes of Verbs and Subjects

KOU MUKAINAKA <sup>†</sup>

This paper presents a model to analyze the coherence relation within complex sentences by using the attributes of verbs and subjects. The relations between subordinate and main clauses can not be understood only by the connectives that link them. The coherence relations by connectives are often ambiguous. For example, a Japanese conjunctive particle “te” expresses coherence relations, such as sequential, method, manner, reason, purpose and parallel. The coherence relations by connectives depend on the semantic types of the predicates in subordinate and main clauses, and the combination of them. It is presumed that there are the patterns of coherence relations in the relations between subordinate and main clauses, just like there are the patterns of thematic roles to express semantic relations between verbs and nouns. The pattern of coherence relation uses the attributes of verbs and subjects in a subordinate and a main clause, just like the pattern of thematic roles uses a verb and the semantic types of nouns. The model infers coherence relations by using the attributes of verbs and subjects in subordinate and main clauses. It uses volition, semantic types, idiomatic expressions, mood/aspect/voice, as the attributes of verbs, and uses whether the subject in a subordinate clause is identical with the one in the main clause or not, whether the subject is unanimate or not as the attributes of subjects. The model is evaluated and the result shows that 95% of the text taken from science and technical documents can be analyzed successfully.

**Keywords:** *Coherence relation, Disambiguation, Attribute of verb, Attribute of subject, Coherence relation pattern*

## 1 はじめに

連接関係の関係の意味は、接続詞、助詞等により一意に決まるものもあるが、一般的には曖昧性を含む場合が多い。一般的には、複文の連接関係の关系的意味は、従属節や主節の表している事象の意味、およびそれらの事象の相互関係によって決まってくる。しかし、各々の単文の意味とそれらの間の関係を理解するためには広範囲の知識が必要になる。それらの背景知識を記述して、談話理解に利用する研究 (Zadrozny and Jensen 1991; Dahlgren 1988, など) も行われているが、現状では、非常に範囲を限定したモデルでなければ実現できない。従って、連接関係を解析するためには、少なくともどのような知識が必要になり、それをを用いてどのように解析するのが問題になる。

シテ型接続に関する研究 (仁田 1995) では、助詞「て」による連接関係を解析し、「時間的継起」のほかに「方法」、「付帯状態」、「理由」、「目的」、「並列」などの意味があることを述べている。これらの关系的意味は、動詞の意志性、意味分類、アスペクト、慣用的な表現、同一主体、無生物主体などによって決まることを解析している。

しかし、動詞の意志性自体が、動詞の語義や文脈によって決まる場合が多い。また、主体が省略されていることも多い。さらに、「て」以外の接続の表現に対して、同じ属性で識別できるかどうか不明である。

表層表現中の情報に基づいて、文章構造を理解しようとする研究 (黒橋・長尾 1994) では、種々の手掛かり表現、同一／同種の語／句の出現、2文間の類似性を利用することによって連接関係を推定している。しかし、手掛かり表現に多義のある時は、ある程度の意味情報を用いる必要がある。

日本語マニュアル文においてアスペクトにより省略された主語を推定する研究 (中川・森 1995) や、知覚思考、心理、言語活動、感情、動きなど述語の意味分類を用いて、「ので」順接複文における意味解析を行う研究 (木村, 西沢, 中川 1996) などがあり、アスペクトや動詞の意味分類が連接関係の意味解析に有効なことが分かる。しかし、連接関係全般について、動詞と主体のどのような属性を用いて、どの程度まで解析できるかが分からない。

本論文では、「て」以外の曖昧性の多い接続の表現についても、その意味を識別するために必要な属性を調べ、曖昧性を解消するモデルを作成した。<sup>1</sup> 動詞の意志性については、予め単文で動詞の格パターンを適用して解析して、できるだけ曖昧性を無くすようにした。省略された主体については、技術論文、解説書、マニュアルなどの技術文書を前提にして、必要な属性を復元するようにした。

† 江戸川大学社会学部 環境情報学科, Department of Environmental Information, Edogawa University

## 2 接続関係の曖昧性

接続詞、助詞等の接続の表現には曖昧性がある。特に、「て」、「ため」、「が」、「と」などで表現される従属節の関係の意味は様々である。これは、「から」、「ので」などと違って、これらの接続の表現自体が明確な固有の意味を有していないからである。従って、これらの接続の表現では、接続関係の関係の意味が、従属節や主節の表している事象の意味、およびそれらの事象の相互関係によって決まってくる。

「～して」形式で表される接続関係の関係の意味には、次の幾つかの例で示すように、「時間的継起」、「方法」、「付帯状態」、「原因」、「目的」、「並列」などがある。

〔例文 1〕 それぞれのセグメントに TCP ヘッダを加えて、相手のプロトコルモジュールに送っています。(時間的継起)

〔例文 2〕 ネットワークを利用して、常に最新のデータを取り出すことができます。(方法)

〔例文 3〕 そこで女中が鍵を持って、私を待っていた。(付帯状態)

〔例文 4〕 この部屋は静かで、よく眠れる。(原因)

〔例文 5〕 私がだれかにこの暗号を伝達する前に詳細を知ろうとして、私をつけ狙った。(目的)

〔例文 6〕 交通は混乱して、人心は陰悪である。(並列)

「～と」形式で表される接続関係の関係の意味には、「時」、「条件」、「原因」などがある。

〔例文 7〕 朝起きると、すぐシャワーを浴びる。(時)

〔例文 8〕 同じユーザ名でもユーザ ID やグループ ID が異なっていると、うまくログインやファイルのコピーができません。(条件)

〔例文 9〕 窓を開けると、寒い風が入った。(原因)

「～ため」形式で表される接続関係の関係の意味には、「原因」と「目的」がある。

〔例文 10〕 AとCは、異なったネットワークにあるため、データリンク層のプロトコルも違います。(原因)

〔例文 12〕 異なる地域、都市、国の間を結ぶため、自分達で勝手に通信回線の敷設はできません。(目的)

「～が」形式で表される接続関係の関係の意味には、「逆接」、「対比」、「前置き」などがある。

〔例文 13〕 何度も説明しましたが、あの人は分からなかった。(逆接)

〔例文 14〕 兄は勤勉だが、弟はずぼらだ。(対比)

〔例文 15〕 ちょっと伺いますが、駅はどこですか。(前置き)

## 3 動詞と主体の属性と接続関係の関係の意味

複文中の接続関係は、従属節や主節の述語の表している事象の意味タイプ、およびその組み合わせによって決まってくる。ここで、事象の意味タイプは、ある主体が行う動作または状態の

分類、意志性などを表す(仁田 1995)。例えば、生物主体の姿勢変化、携帯、心的状態、使用、作成、助言などである。従って、事象の意味タイプは、動詞と名詞の属性を用いて表すことができると思うことができる。従って、動詞と名詞の意味的關係を表すために、動詞と名詞の意味分類を用いた格パターンがあると同様に、従属節と主節の連接關係にも、動詞と名詞の属性を用いた連接關係パターンが存在すると思うことができる。

本論文では、従属節と主節の、動詞と主体の属性を用いて、連接關係の關係的意味を推定する方法をとった。動詞の属性として、意志性、意味分類、慣用的表現、ムード・アспект・ヴォイス、主体の属性として、主節と従属節の主体が同一かどうか、無生物主体かどうかを採用した。

次に、各々の属性によって、連接關係の關係的意味がどのように決まるかを、いくつかの例で示す。

### 3.1 動詞の意志性と主体の同一性、無生物性

従属節と主節の動詞の意志性および主体が同一かどうか、無生物主体かどうかの組み合わせによって、連接關係の關係的意味が次のように影響を受ける。ここで、主体は、動作主、経験者などを含む概念である(仁田 1995)。動詞の意志性は、人間の意志的な行為を表し、誘い掛けや命令の意味を表す。生物/無生物は、生命の有る/無しではなく、情意、特に自由意志による行動可能なものを表す。

「～して」形式接続で、従属節と主節が共に意志動詞で形成され、両者の主体が同一の時は、「時間的継起」を表すことが多い。これは、同一主体による制御可能な動きは、通常、継起的に引き起こされるためである。本論文のモデルで用いた用例では、このパターンに属する連接關係の94%が「時間的継起」を表した。ただし、後に述べる従属節の意味分類から「方法」、「付帯状態」と識別されるものは除いてある。用例の内容については5章で述べる。

〔例文16〕 ユーザは、IDとパスワードを指定して、OKボタンをクリックします。

「～して」形式接続で、従属節と主節の主体が無生物で動詞が共に無意志動詞の時は、「時間的継起」を表すことが多い。自然界における無生物的な2現象が、時間的継起の下に生じることとはよくある。この場合がそれに相当する。用例では、従属節の意味分類から「原因」と識別されるものを除くと、このパターンに属する連接關係の85%が「時間的継起」を表した。

〔例文17〕 ログインが成立して、プロンプトが戻ってきます。

「～して」形式接続で、従属節と主節が共に無意志動詞で従属節が無生物主体、主節が生物主体の時は、「原因」を表すことが多い。これは、主たる事象が人間に関するものでありながら、従属節で人間の意志で制御できない事象が生じたためである。一般に技術文書ではこのパターンは少なく、用例でも2例しかなかったが、2例の連接關係はいずれも「原因」を表した。

〔例文18〕 彼は、車が故障して、遅れた。

「～ため」形式接続で、従属節と主節が共に意志動詞で生物主体の時は、「目的」を表すことが多い。これは、同一主体が、従属節の事象を達成する目的で、主たる事象を行うためである。用例ではこのパターンに適合する接続関係は、全て「目的」を表した。

〔例文 19〕 ユーザは、FDDI で UNIX ワークステーションを接続するために、通常 FDDI の通信用ボードを購入しなければなりません。

「～ため」形式接続で、従属節が無意志動詞で、主節も無意志動詞の時は「原因」を表すことが多い。これは、従属節で人間の意志で制御できない事象が生じ、それからある事象が起こったときは、その事象の原因と解釈されるためである。

無意志動詞には、心的作用を表す動詞や、状態を表す動詞のほか、意志動詞に「される」が後接した受動態や、意志動詞に「ている」、「である」が後接した「単純状態」や「変化状態の維持」のアスペクトなどを含む。このような場合を含めると、用例では、このパターンに適合する接続関係はすべて「原因」を表した。

〔例文 20〕 FDDI はトークンパッシング方式を採用しているため、ネットワークトラフィックによる性能の低下がない。

「～と」形式接続で、従属節と主節が共に意志動詞で同一主体の時は、「時」を表すことが多い。「～と」形式接続は一般的には「条件」を表すことが多いが、同一主体による制御可能な動きの場合には、「その時」または「～してすぐ」という意味になりやすい。主節のヴォイスが「可能」を表す時には「条件」を表すことが多いので、この場合を除くと、用例ではこのパターンに適合する接続関係は全て「時」を表した。

〔例文 21〕 彼女は、部屋に入ると、窓を開けた。

### 3.2 動詞の意味分類

動詞の意味分類によっても、接続関係の关系的意味が影響を受けることがある。次に、幾つかの例を挙げる。

「～して」形式接続で従属節の動詞が、姿勢変化、着脱、携帯、心的状態などの意味分類であるときは「付帯状態」を表すことが多い。「付帯状態」とは、従属節と主節の事象が時間的に同存し、同一主体で、従属節で主節の事象の実現のされ方を表しているものである。

〔例文 22〕 わたしは汗で湿った服をそのまま着て、また寮に出かけていった。

「～して」形式接続で、従属節の動詞が、使用、作成、助言などの意味分類であるときは、「方法」を表すことが多い。つまり、従属節の事象が、主節の事象を実現するための方法的要因になっている場合である。技術文書では、「～を利用して」、「～を作成して」などの表現が多いので、接続関係の关系的意味にも「方法」がかなりある。意味分類の定義は、基本的には分類語彙表に因った。分類語彙表の意味分類に用例を適用して、正の用例のみを含む場合は分類語彙表の分類で定義したが、負の用例を含む場合は正の用例の単語をそのまま定義に加えた。

〔例文 2 3〕 そのケーブルを使って、データを送る必要があります。

「～が」形式接続で、主節と従属節の動詞又は形容詞の意味分類が、反意語の時は「対比」を表すことが多い。用例では、2 例だけであったが、2 例とも「対比」を表した。

〔例文 2 4〕 昨日まで寒かったが、今日から急に暖かくなった。

### 3.3 ムード・アспект・ヴォイス

ムード・アспект・ヴォイスによっても接続関係の關係的意味が変わってくる。前述のように、「される」が後接した受動態や、「ている」、「である」が後接した「單純状態」や「変化状態の維持」のアспектなどを含む節は無意志的に解釈されるが、そのほかにも次に示すような幾つかの例がある。

「～して」形式接続で、従属節が「～（よ）うとして」といった將然相の形をとる場合は、「目的」を表すことが多い。主節の事象を引き起こす計画を、従属節で述べているためである。

〔例文 2 5〕 体を鍛えようとして、毎日ジョギングをやっている。

「～が」形式接続で、主節が疑問文の時は、「前置き」を表すことが多い。一般的には、「～が」は「逆接」または「対比」を表すことが多いが、主節が疑問文の時は、質問に対する前提条件を表すことが多いためである。用例では、このパターンに属する接続関係は全て「前置き」を表した。

〔例文 2 6〕 ここに鍵が置いてありますが、誰のですか。

「～と」形式接続で、従属節が生物主体で「ている」、「かける」、「はじめる」の場合は、「時」を表すことが多い。「～と」は一般的には、「条件」を表すことが多いが、従属節が生物主体で動作の「進行」または「開始」を表すアспектの場合は、「その時」または「～してすぐ」の意味になることが多いためである。用例では、このパターンは少なかったが、適合する接続関係は全て「時」を表した。

〔例文 2 7〕 食事をしていると、急にグラッと揺れた。

### 3.4 従属節が慣用句的になっているもの

従属節が慣用句化して副詞的に用いられる場合がある。この場合は、接続の表現もそれぞれの慣用句に対応した關係的意味を持つ。

「～して」形式接続で、従属節が「体力をふり絞って」、「先を争って」、「まとまって」、「だまって」などの表現をとるときは、「付帶状態」を表す。

〔例文 2 8〕 体力を振り絞って、走った。

## 4 複文の接続関係解析モデル

前述のように複文の接続関係は、動詞の意志性、意味分類、ムード・アスペクト・ヴォイス、慣用表現、主体の同一性、無生物主体かどうかなどによって決まる場合が多い。これらの情報により、接続関係を解析することができる。しかし、動詞の意志性自体に曖昧性がある。また、主体が省略されていることも多い。従って、接続関係を解析する前に、これらの情報を解析しておく必要がある。複文の接続関係の解析モデルは図1に示すように、動詞と主体の属性を解析し、それらの属性を用いて、接続関係の解析を行う。

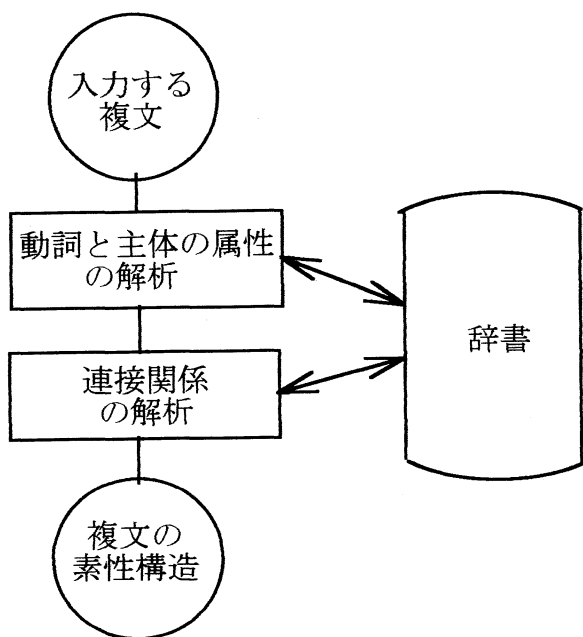


図1 接続関係の解析モデル

本論文のモデルは連用修飾節の解析を対象としたので、「の」、「こと」などによる連体修飾節は、入力文から省いてある。接続関係パターン、格パターン、名詞の属性などを HPSG(Pollard and Sag 1987, 1993, など) の素性構造に似た形式で表し、辞書に登録した。これらの辞書の情報を用いて、動詞と主体の属性の解析、接続関係の解析を行い、複文の素性構造に相当する出力を生成する。解析は、簡単な HPSG パーザを prolog で作成して行った。このパーザは、格パターンの解析と接続関係パターンの解析の機能だけを持ったもので、この目的のために作成した。

辞書は、IPAL 辞書 (情報処理振興事業協会 1987, 1990, など) と分類語彙表 (国立国語研究所 1989) に基づき作成した。動詞の格パターン、動詞の意志性は IPAL 辞書のものを採用した。動詞の意味分類は分類語彙表の分類を用いた。IPAL 辞書の格パターンで用いている名詞の意味分類は、比較的粗い分類になっている。従って、動詞の意味分類の方が名詞の意味分類よりも詳しくなっている。

名詞の生物／無生物の区分については、意味分類が「人間」と「組織」の場合を「生物」とし、それ以外を「無生物」とした。用例は、主としてネットワークプログラムの解説書からとったため、プログラムの機能説明が多数含まれていた。その中で、プログラムが生物と同様の行動をするので、プログラムも生物に分類した。コンピュータ用語で IPAL 辞書にも分類語彙表にも無い用語が多数出てきたが、同じ分類の用語に準じて定義し辞書に追加した。

#### 4.1 動詞と主体の属性の解析

複数の語義のある動詞は、語義を確定しないと、意志性を確定できない。そのため、動詞の格パターンを適用して、語義を確定するようにした。図 2 に、辞書における動詞の格パターンの記載例を示す。図 2 では HPSG の形式で記載しているが、実際の辞書では prolog の形式に変換して格納している。この格パターンを用いて入力文を解析する過程で、適合したパターンの語義を選択し、意志性を確定した。

しかし、意志動詞でも、特定の語義で無意志用法のあるものがある。

【例文 2 9】 B29 は工場に爆弾を落とした。(意志用法)

【例文 3 0】 彼女はお皿を落として割ってしまった。(無意志用法)

ただし、技術文書を考えた場合、意志動詞の無意志用法は比較的少ない。実際に用例を調べた結果、意志動詞の無意志用法は、非常に少く 2% であったので無視することにした。従って、IPAL 辞書の動詞の意志性で無意志用法の有るものは、意志動詞として分類した。

主体が省略されている時には、マニュアル、技術論文などの技術文書であることを前提にして、比較的単純な方式により、生物主体か無生物主体か、同一主体か異主体かを推定するようにした。

一般的に、著者や読者が主題になっているときは、先行文脈から推定可能であり、特に強調する必要があるときなど、特別の場合以外は、省略されるのが普通である。例えば、マニュアル類では、装置の開発者が、利用者に説明することを前提にして書かれているので、開発者や利用者が省略されることが多い。技術文書でも、開発者や研究者又は読者が省略される。

文全体の主題と従属節の主題が同じ時は、従属節の主題は省略される。ただし、従属節に主題でない主体があるにもかかわらず、主節の主体が省略されることは有り得る。この場合は、省略された主体は、先行文脈の主題であり、かつ、文全体が主題寄りの視点で記述されている可能性が高い (久野 1978)。技術文書では著者又は読者寄りの視点から書かれているのが普通で



## (a) 動詞の記載例

<i>PHON</i>	開ける													
<i>SYN LOC CAT</i>	<table> <tr> <td><i>HEAD</i></td> <td> <table> <tr> <td><i>MAJ</i></td> <td><i>V</i></td> </tr> <tr> <td><i>VFORM</i></td> <td><i>DICTIONARY - FORM</i></td> </tr> <tr> <td><i>ADJUNCTS</i></td> <td>{<i>[LOC HEAD MAJ ADV]</i>}</td> </tr> <tr> <td><i>LEX</i></td> <td>+</td> </tr> </table> </td> </tr> <tr> <td><i>SUBCAT</i></td> <td> <math>\left\langle \begin{array}{l} PP[NOM][1][+ANIMATE, HUMAN], \\ PP[ACC][2][-ANIMATE, PRODUCT] \end{array} \right\rangle</math> </td> </tr> </table>	<i>HEAD</i>	<table> <tr> <td><i>MAJ</i></td> <td><i>V</i></td> </tr> <tr> <td><i>VFORM</i></td> <td><i>DICTIONARY - FORM</i></td> </tr> <tr> <td><i>ADJUNCTS</i></td> <td>{<i>[LOC HEAD MAJ ADV]</i>}</td> </tr> <tr> <td><i>LEX</i></td> <td>+</td> </tr> </table>	<i>MAJ</i>	<i>V</i>	<i>VFORM</i>	<i>DICTIONARY - FORM</i>	<i>ADJUNCTS</i>	{ <i>[LOC HEAD MAJ ADV]</i> }	<i>LEX</i>	+	<i>SUBCAT</i>	$\left\langle \begin{array}{l} PP[NOM][1][+ANIMATE, HUMAN], \\ PP[ACC][2][-ANIMATE, PRODUCT] \end{array} \right\rangle$	
<i>HEAD</i>	<table> <tr> <td><i>MAJ</i></td> <td><i>V</i></td> </tr> <tr> <td><i>VFORM</i></td> <td><i>DICTIONARY - FORM</i></td> </tr> <tr> <td><i>ADJUNCTS</i></td> <td>{<i>[LOC HEAD MAJ ADV]</i>}</td> </tr> <tr> <td><i>LEX</i></td> <td>+</td> </tr> </table>	<i>MAJ</i>	<i>V</i>	<i>VFORM</i>	<i>DICTIONARY - FORM</i>	<i>ADJUNCTS</i>	{ <i>[LOC HEAD MAJ ADV]</i> }	<i>LEX</i>	+					
<i>MAJ</i>	<i>V</i>													
<i>VFORM</i>	<i>DICTIONARY - FORM</i>													
<i>ADJUNCTS</i>	{ <i>[LOC HEAD MAJ ADV]</i> }													
<i>LEX</i>	+													
<i>SUBCAT</i>	$\left\langle \begin{array}{l} PP[NOM][1][+ANIMATE, HUMAN], \\ PP[ACC][2][-ANIMATE, PRODUCT] \end{array} \right\rangle$													
<i>SEM CONT</i>	<table> <tr> <td><i>RELN</i></td> <td><i>AKERU</i></td> </tr> <tr> <td><i>OPENER</i></td> <td>[1]</td> </tr> <tr> <td><i>OPEND</i></td> <td>[2]</td> </tr> <tr> <td><i>VOLITION</i></td> <td>+</td> </tr> <tr> <td><i>SEM - TYPE</i></td> <td>開・閉</td> </tr> </table>	<i>RELN</i>	<i>AKERU</i>	<i>OPENER</i>	[1]	<i>OPEND</i>	[2]	<i>VOLITION</i>	+	<i>SEM - TYPE</i>	開・閉			
<i>RELN</i>	<i>AKERU</i>													
<i>OPENER</i>	[1]													
<i>OPEND</i>	[2]													
<i>VOLITION</i>	+													
<i>SEM - TYPE</i>	開・閉													

注) *VP[DICTIONARY - FORM][+VOLITION]* と略記する

(b) 名詞句 *PP[ACC][-ANIMATE, PRODUCT]* の内容

<i>PHON</i>	箱を											
<i>SYN LOC CAT</i>	<table> <tr> <td><i>HEAD</i></td> <td> <table> <tr> <td><i>MAJ</i></td> <td><i>P</i></td> </tr> <tr> <td><i>PFORM</i></td> <td><i>WO</i></td> </tr> <tr> <td><i>CASE</i></td> <td><i>ACC</i></td> </tr> </table> </td> </tr> <tr> <td><i>SUBCAT</i></td> <td><math>\langle NP[-ANIMATE, PRODUCT] \rangle</math></td> </tr> </table>	<i>HEAD</i>	<table> <tr> <td><i>MAJ</i></td> <td><i>P</i></td> </tr> <tr> <td><i>PFORM</i></td> <td><i>WO</i></td> </tr> <tr> <td><i>CASE</i></td> <td><i>ACC</i></td> </tr> </table>	<i>MAJ</i>	<i>P</i>	<i>PFORM</i>	<i>WO</i>	<i>CASE</i>	<i>ACC</i>	<i>SUBCAT</i>	$\langle NP[-ANIMATE, PRODUCT] \rangle$	
<i>HEAD</i>	<table> <tr> <td><i>MAJ</i></td> <td><i>P</i></td> </tr> <tr> <td><i>PFORM</i></td> <td><i>WO</i></td> </tr> <tr> <td><i>CASE</i></td> <td><i>ACC</i></td> </tr> </table>	<i>MAJ</i>	<i>P</i>	<i>PFORM</i>	<i>WO</i>	<i>CASE</i>	<i>ACC</i>					
<i>MAJ</i>	<i>P</i>											
<i>PFORM</i>	<i>WO</i>											
<i>CASE</i>	<i>ACC</i>											
<i>SUBCAT</i>	$\langle NP[-ANIMATE, PRODUCT] \rangle$											

(c) 名詞 *NP[-ANIMATE, PRODUCT]* の内容

<i>PHON</i>	箱					
<i>SYN LOC CAT</i>	<table> <tr> <td><i>HEAD MAJ</i></td> <td><i>N</i></td> </tr> <tr> <td><i>SUBCAT</i></td> <td><math>\langle \rangle</math></td> </tr> </table>	<i>HEAD MAJ</i>	<i>N</i>	<i>SUBCAT</i>	$\langle \rangle$	
<i>HEAD MAJ</i>	<i>N</i>					
<i>SUBCAT</i>	$\langle \rangle$					
<i>SEM LOC CONT</i>	<table> <tr> <td><i>ANIMATE</i></td> <td>-</td> </tr> <tr> <td><i>SEM - TYPE</i></td> <td><i>PRODUCT</i></td> </tr> </table>	<i>ANIMATE</i>	-	<i>SEM - TYPE</i>	<i>PRODUCT</i>	
<i>ANIMATE</i>	-					
<i>SEM - TYPE</i>	<i>PRODUCT</i>					

図 2 辞書における動詞の格パターンの記載例

あるから、著者又は読者が省略されている場合が多い。

省略された主体を埋める候補が意味的に矛盾しないかどうかの検証は、動詞の要求する主体が生物か無生物かによった。

接続関係の解析に必要な情報は、同一主体か、無生物主体かだけであり、HPSG パーザは格パターン、接続関係パターンの順序で縦形探索を用いて解析しているので、次のような簡単な処理方法により省略された主体を推定した。

- (1) 格パターンを適用するとき、生物主体のパターンを優先する。
- (2) 最終的に動詞の格パターンの SUBCAT に残った項を、省略を復元する候補とする。
- (3) 接続関係パターンを適用するとき、同一主体のパターンを優先する。
- (4) 従属節が無生物主体で、主節の主体が省略されている時だけ特別扱いし、生物/無生物の異主体のパターンを優先する。

(5) 優先するパターンが無いときは、次に確率の高いパターンを適用する。

1 は、著者や読者が省略されている可能性が高い事を表す規則であり、3 は、次の4 の場合を除いて、文全体の主題と従属節の主題が同じ場合、および従属節の主体と主節の主体が同じ場合が多い事を表す規則である。4 は、従属節が無生物主体で、主節の主体が著者または読者の場合に相当する。

従属節の動詞の意味分類によって、接続関係の关系的意味が決まってくることもある。従って、動詞が特定の意味分類に属しているかどうかを、解析する必要がある。このため、動詞の意味分類によって決まる接続の表現に対しては、辞書の接続関係パターンの記載に、その关系的意味が要求する動詞の意味分類を記載して、動詞の意味分類と単一化するようにした。動詞の意味分類は、前述のように接続関係の关系的意味を識別する必要上、名詞の意味分類より詳細になっている。

従属節が慣用句化して副詞的に用いられる場合は、慣用句として辞書に記載しておき、入力文を解析する時に優先的に選択する。

上記のように HPSG パーザの解析過程で、格パターンを適用することにより、動詞と主体の属性を求め、引き続き接続関係パターンを適用することにより、接続関係を解析する。

## 4.2 接続関係の解析

接続の表現は、図3に示すような形式で辞書に記載される。図3には、「～して」形式接続で、従属節と主節が共に意志動詞、主体が同一で、「時間的継起」を表す場合の接続関係パターンを示す。このような接続関係パターンが、接続の表現別、接続関係の关系的意味別に存在する。1つの关系的意味を複数のパターンで表すことも有り得る。

$$\left[ \begin{array}{l} PHON \\ SYN \\ SEM|LOC|CONT \end{array} \left[ \begin{array}{l} \text{て} \\ LOC|CAT \left[ \begin{array}{l} HEAD \left[ \begin{array}{l} MAJ \\ JFORM \end{array} \right] \left[ \begin{array}{l} ADV \\ TE \end{array} \right] \\ SUBCAT \left\langle \begin{array}{l} VP[ADVERBIAL - FORM][1][+VOLITION], \\ SUBCAT(PP[NOM][3][+ANIMATE]) \end{array} \right\rangle \end{array} \right] \\ BIND|CAT|MAIN \left\langle \begin{array}{l} VP[2][+VOLITION], \\ SUBCAT(PP[NOM][3][+ANIMATE]) \end{array} \right\rangle \end{array} \right] \right] \right]$$

図3 辞書における接続関係パターンの記載例

接続関係パターンで使用している素性は、表 1 の通りである。ただし、一般的な HPSG の素性は省いてある。

表 1 接続関係パターンの素性

素性	値	意味
VOLITION	+, -	意志性
ANIMATE	+, -	生物/無生物
SEM-TYPE	{ 使用, 製造, 教育 }, ...	意味分類
VOICE	PASSIVE, ACTIVE, CAUSATIVE, POSSIBLE	ヴォイス (受動/能動/使役/可能)
ASPECT	{ TEIRU, TEARU }, ...	アスペクト
MOOD	QUESTION, VOLITION	ムード (疑問/意志)
IDIOM-TYPE	TE-SEQ, GA-PRE	慣用句の分類

特定の接続の表現の接続関係パターンで、従属節の要求する素性は SUBCAT で表され、主節の要求する素性は MAIN で表される。解析の過程で単一化に成功したパターンが選択され、適合する接続関係の关系的意味が選択されることになる。

## 5 接続関係解析モデルの評価結果

本論文の接続関係の解析モデルを、実際の技術文書に適用して評価した。評価結果は、表 2 に示す通りである。

まず、接続の表現として最も多く使われる 20 の接続の表現を選択した。表 2 には、この 20 の表現に、評価した例文に出てきた 4 つの表現を追加して挙げてある。複文中の単文と単文を接続する表現を中心として選択し、複文と複文を接続する表現は除いた。また、本論文では、前述のように連用修飾節を解析の対象としたので、「の」、「こと」などによる連体修飾節に対する接続の表現は省いてある。

上記 20 の接続の表現に対する接続関係パターンを解析した。解析のための例文としては、複文に関する論文集 (仁田 1995) から 145 文、日本語教育の参考書 (横林・下村 1988; 北条 1992, など) から 323 文、ネットワークの解説書 (金内・今安 1993) の前半から 312 文を選択した。合計 780 文の接続関係に適合する 52 の接続関係パターンを抽出した。接続関係パターンは図 3 に示すような構成をしており、表 1 に示す素性の有／無がパターン毎に異なる。各々の接続の表現毎の关系的意味に対するパターン数は表 2 に示す通りである。

この接続関係パターンを、ネットワークの解説書の後半からとった別の 238 文に適用した結果が、表 2 の接続関係モデルの評価結果である。用意した 20 の接続の表現に含まれないものが 4 例、正しく解釈できなかったものが 8 例あった。合計して 5% の文は正しく解釈できなかったが、残りの 95% は正しく解析することができた。正しく解析できたかどうかの判断の基準は、

表 2 接続関係解析モデルの評価結果

接続の表現	関係の意味	パターン数	文数	正しく解析できた文	正しく解析できなかった文
あいだ	時	0	1		1
うえで	目的	1	2	2	
が	逆説	1	12	12	
	前置き	2	4	4	
	対比	1	3	1	2
から	原因	1	3	3	
	理由	1	1	1	
さいに	時	1	4	4	
し	並列	1	6	6	
	理由	1	0		
ため(に)	目的	2	13	13	
	原因	2	9	8	1
たら(ば)	条件	1	1	1	
	理由	1	0		
	時	1	0		
たり	並列	1	4	4	
て/で	時間的継起	3	14	10	4
	方法	3	8	8	
	付帯状態	1	6	6	
	原因	2	4	4	
	目的	1	0		
	並列	1	1	1	
ても	逆接条件	1	10	10	
と	条件	2	13	13	
	時	2	2	2	
	原因	1	0		
といえ	題材	0	1		1
ば	時	1	9	9	
とき	時	0	1		1
とともに	同時動作	1	1	1	
ながら	逆接	1	0		
なら(ば)	条件	1	0		
ので	原因	1	12	12	
	理由	1	14	14	
のに	逆接条件	1	2	2	
ば	条件	1	22	22	
	並列	1	0		
ばあい	条件	1	14	14	
ように	対比	2	3	3	
	推量	1	1	1	
より	対比	0	1		1
連用中止	並列	2	28	28	
	時間的継起	2	8	7	1
合計		52	238	226	12
比率 (%)			100	95	5

「～して」形式接続については複文に関する論文集(仁田 1995), その他については日本語教育の参考書(横林・下村 1988; 北条 1992, など)の例文に因った。似た例文を探して, その例文の接続関係の意味を採用した。接続関係の意味が正しいかどうかの判断の結果は, 3名の語学の研究者に見てもらい, 異議の出たものは修正した。

## 6 むすび

本論文では, 従属節と主節の, 動詞と主体の属性を用いて, 接続関係の关系的意味を解析するシステムを作成した。動詞と名詞の意味的關係を表すために, 動詞と名詞の意味分類を用いた格パターンがあると同様に, 従属節と主節の接続関係にも, 各々の節を構成する動詞と主体の属性を用いた接続関係パターンが存在すると思われることができる。動詞の属性として, 動詞の意志性, 意味分類, 慣用的表現, ムード・アスペクト・ヴォイス, 主体の属性として, 主節と従属節の主体が同一かどうか, 無生物主体かどうかを採用した。

このシステムを, 実際の技術文書に適用して評価した結果, 95%の正しい解析結果を得ることができた。

今回は, 接続関係パターンを手作業で抽出したが, 確率モデルを採用し, パターンの属性の組み合わせを学習するシステムに拡張する事が今後の研究課題である。

## 参考文献

- Dahlgren, K. (1988). *Naive Semantics for Natural Language Understanding*. Kluwer Academic Publishers.
- 北条淳子 (1992). “複文文型.” 談話の研究と教育, pp. 7-112. 国立国語研究所.
- 仁田義雄 (1995). 複文の研究 (上) : シテ形接続をめぐる. くろしお出版.
- 情報処理振興事業協会 (1987). 計算機用日本語基本動詞辞書 IPAL (Basic Verbs). 情報処理振興事業協会技術センター.
- 情報処理振興事業協会 (1990). 計算機用日本語基本形容詞辞書 IPAL (Basic Adjectives). 情報処理振興事業協会技術センター.
- 金内典充・今安正和 (1993). UNIX ネットワークプログラミング. オーム社.
- 木村啓一, 西沢信一郎, 中川裕志 (1996). “述語の意味を用いた日本語順接複文の理解システム.” 情報処理学会自然言語処理研究会 96-NL-111, pp. 47-54.
- 国立国語研究所 (1989). 分類語彙表. 秀英出版.
- 久野日章 (1978). 談話の文法. 大修館書店.
- 黒橋慎夫・長尾真 (1994). “表層表現中の情報に基づく文章構造の自動抽出.” 自然言語処理, 1 (1), 3-20.

中川裕志・森辰則 (1995). “日本語マニュアル文におけるテイル, テアル, テオク, テミル, テシマウの語用論.” 自然言語処理, **2** (4), 19-35.

Pollard, C. and Sag, I. (1987). *Information-Based Syntax and Semantics Vol. 1*. CSLI Stanford University.

Pollard, C. and Sag, I. (1993). *Head-Driven Phrase Structure Grammar*. The University of Chicago Press.

横林宙世・下村彰子 (1988). 接続の表現. 荒竹出版.

Zadrozny, W. and Jensen, K. (1991). “Semantics of Paragraphs.” *Computational Linguistics*, **17** (2), 171-209.

## 略歴

**向仲 景貞:** 1953 九州大学工学部電気工学科卒業. 同年, 日本電気(株)入社. 基本ソフトウェア開発に従事. 平成5年より金沢経済大学教授. 平成9年より江戸川大学教授, 現在に至る. 自然言語理解, エキスパートシステムの研究に従事. 情報処理学会, 言語処理学会, 人工知能学会, ACL, ACM, IEEE 各会員.

(1996 年 4 月 24 日 受付)

(1997 年 5 月 20 日 再受付)

(1997 年 7 月 18 日 採録)